

BA



The discovery of new G-protein coupled receptors, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders, and viral infections, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of G-protein coupled receptors.

TECH CENTER 1600/2900

JUL 15 2002

RECEIVED

### SUMMARY OF THE INVENTION

The invention features purified polypeptides, G-protein coupled receptors, referred to collectively as "GCREC" and individually as "GCREC-1," "GCREC-2," "GCREC-3," "GCREC-4," "GCREC-5," "GCREC-6," "GCREC-7," "GCREC-8," "GCREC-9," "GCREC-10," "GCREC-11," "GCREC-12," "GCREC-13," "GCREC-14," "GCREC-15," and "GCREC-16." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-16.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-16. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:17-32.

The invention additionally provides G-protein coupled receptors that are involved in olfactory and/or taste sensation. The invention further provides polynucleotide sequences that encode said G-protein coupled receptors.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting

of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino  
5 acid sequence selected from the group consisting of SEQ ID NO:1-16. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting  
10 of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. The method comprises a)  
15 culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence  
20 selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ  
25 ID NO:1-16.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32,  
30 c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of  
35 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ

5 ID NO:17-32, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe  
10 comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and,  
15 optionally, if present, the amount thereof.  
20

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. The invention additionally  
25 provides a method of treating a disease or condition associated with decreased expression of functional GCREC, comprising administering to a patient in need of such treatment the composition.  
30

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected  
35

naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-16. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides methods of using G-protein coupled receptors of the invention involved in olfactory and/or taste sensation, biologically active fragments thereof (including those having receptor activity), and amino acid sequences having at least 90% sequence identity therewith, to identify compounds that agonize or antagonize the foregoing receptor polypeptides. These compounds are useful for modulating, blocking and/or mimicking specific tastes and/or odors.

The present invention also relates to the use of olfactory and/or taste receptors of the invention, biologically active fragments thereof (including those having receptor activity), and polypeptides having at least 90% sequence identity therewith, in combination with one or more other olfactory and/or taste receptor polypeptides, to identify a compound or plurality of compounds that modulate, mimic, and/or block a specific olfactory and/or taste sensation.

The invention also relates to cells that express an olfactory or taste receptor polypeptide of the invention, a biologically active fragment thereof (including those having receptor activity), or a polypeptide having at least 90% sequence identity therewith, and the use of such cells in cell-based screens to identify molecules that modulate, mimic, and/or block specific olfactory or taste sensations.

Still further, the invention relates to a cell that co-expresses at least one olfactory or taste G-protein coupled receptor polypeptide of the invention, and a G-protein, and optionally one or more other olfactory and/or taste G-protein coupled receptor polypeptides, and the use of such a cell in screens to identify molecules that modulate, mimic, and/or block specific olfactory and/or taste sensations.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b)

hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a  
5 polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising  
10 a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:17-32, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment  
15 of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

20

#### BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank  
25 homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

30 Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of  
35 the cDNA libraries shown in Table 5.

assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of GCREC or the polynucleotide encoding GCREC which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:17-32 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:17-32, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:17-32 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:17-32 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:17-32 and the region of SEQ ID NO:17-32 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-16 is encoded by a fragment of SEQ ID NO:17-32. A fragment of SEQ ID NO:1-16 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-16. For example, a fragment of SEQ ID NO:1-16 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-16. The precise length of a fragment of SEQ ID NO:1-16 and the region of SEQ ID NO:1-16 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a

ID NO:9 also contains a rhodopsin family 7-transmembrane receptor domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS and BLAST analyses provide further corroborative evidence that SEQ ID NO:9 is a rhodopsin family G-protein coupled receptor. In an alternative example, SEQ ID NO:11 is 82% identical to Marmota marmota olfactory receptor (GenBank ID g5901488) as determined by BLAST. (See Table 2.) The BLAST probability score is 1.5e-101. SEQ ID NO:11 also contains a rhodopsin family 7-transmembrane receptor domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:11 is a G-protein coupled receptor. In an alternative example, SEQ ID NO:12 is 64% identical to Homo sapiens olfactory receptor (GenBank ID g2792018) as determined by BLAST. (See Table 2.) The BLAST probability score is 8.4e-99. SEQ ID NO:12 also contains a rhodopsin family 7-transmembrane receptor domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:12 is a G-protein coupled receptor. In an alternative example, SEQ ID NO:15 is 57% identical to chicken olfactory receptor 4 (GenBank ID g1246534) as determined by BLAST. (See Table 2.) The BLAST probability score is 1.1e-91. SEQ ID NO:15 also contains a 7-transmembrane receptor (rhodopsin family) domain as determined by searching the HMM-based PFAM database. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:15 is an olfactory receptor. SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7, SEQ ID NO:10, SEQ ID NO:13-14, and SEQ ID NO:16 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-16 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:17-32 or that distinguish between SEQ ID NO:17-32 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective

5	GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
	GBI	Hand-edited analysis of genomic sequences.
	FL	<i>Stitched or stretched genomic sequences (see Example V).</i>
	5 INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

10 Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

15 The invention also encompasses GCREC variants. A preferred GCREC variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the GCREC amino acid sequence, and which contains at least one functional or structural characteristic of GCREC.

The invention also encompasses polynucleotides which encode GCREC. In a particular 20 embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:17-32, which encodes GCREC. The polynucleotide sequences of SEQ ID NO:17-32, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

25 The invention also encompasses a variant of a polynucleotide sequence encoding GCREC. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding GCREC. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:17-30 32 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ

concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

#### **IX. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:17-32 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-based

Table 4 (cont).

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
21				3406743H1 (PROSTUS08)	103	354
22	6157025CB1	1499	1-81, 248-1499, 850-920, 1265-1499	GNN.g5686520_000054_002.edit 6834181F6 (BRSTNON02) FL6157025_g10567930_000001_g53 53887_1_1	116 826 381	565 1499 1296
23	55012817CB1	2455	1397-1482, 318-704, 1893-2171, 1-1014, 1619-2182	7169316F8 (MCLRNOC01) 4425914H1 (BRAPDIT01) 71691523V1 71691480V1 71691220V1 71690572V1 71691571V1 71691384V1 71688567V1 72488727D1 72489896D1 72487236D1 72492196D1	1 32 1746 1405 1 654 608 1351 2162 841 1 1046 1200	676 288 2401 2082 646 1377 1338 2058 2455 1769 953 2056 2056
24	7475061CB1	2056	1-688, 1207-2056, 1736-1917, 1207-1680	FL7477374_g9930948_000007_6739 496 4021537F6 (BRAXNOT02) GBI.g9967464_000017_000007_000 001.edit	1 1427 1	999 2110 588
25	7477374CB1	999	1-138, 264-999	55017814H1 58013229J1 8042178J1 (OVARTUE01) GNN.g9967464_000015_002 71702678V1 6609077H2 (PLACFEC01) 7720776H1 (THYRDIE01) 7726057J1 (THYRDIE01)	737 2532 2032 729 2160 366 880 1242	1280 3429 2692 2847 2786 877 1291 1974
26	7479890CB1	3429	1-2724, 2850-2900, 1-2300, 2367-2732			

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
17	2536292CB1	BRAINOT18
19	7474823CB1	UCMCL5T01
20	644692CB1	SINTTMR02
21	3837054CB1	DENDTNT01
22	6157025CB1	MONOTXN05
23	55012817CB1	BRAPDIT01
26	7479890CB1	LUNGNON07
30	7478550CB1	BRAUTDR04

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
- 5 73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
- 10 76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
- 15 78. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:17.
79. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:18.
80. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:19.
- 20 81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:20.
82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:21.
- 25 83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:22.
84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:23.
85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:24.
- 30 86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:25.
87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:26.

PI-0236 PCT

```
cagttcctct catatctccc catttgagga cagtggagcca ctggaaagt gtaaaaacaa 1680
aaacagttat ttttgcagtt ttctttcacg catttatagt gctctggata atgccattta 1740
tttttgctga ttacccaact ttccagtatt gctgtgttat catctgtatt tacttatttt 1800
gaatcgtgct taaatcaaact gtaccttcag cacctgcaag tttgcctttt ctttccagga 1860
ggaaaatccc cagcttgctc tccctgggga gtctgagaat tataaccagt ctgtcagaaa 1920
tgtaatcatg ctgtcatttc agagccacag agtatattata aaataaaaac ctttcccacg 1980
gaaaaaaaaa aaa 1993
```

&lt;210&gt; 22

&lt;211&gt; 1499

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6157025CB1

&lt;400&gt; 22

```
aaatgcatgc agagcatgga aatgccccca gctgcctgc tgttgaacaa gaatcctatt 60
tggaaggcag acatgtggcc catctctgta gccatcactg agaaatctgg attttcaagg 120
gcctttctct ctgttgccca ggctggagtt tagcgactca atcatggctc actgactgca 180
gcacgcacct ccggggctca agtgatcctt tcatctcagc ctccctcagta gctgagacta 240
cagggttttc tgggatcagc tgcactcctt agcaaaagta tattggagaa tcaactgaga 300
aagtaactga gacatttcaa tcatttctag gtgtaagaa agaccagatc ccaggaaaat 360
attacggtga cttcccaagt atgcctggcc acaatacctc caggaaattcc tcttgcgatc 420
ctatagtgc accccactta atcagcctct acttcatagt gcttattggc gggctgggtg 480
gtgtcatttc cattcttttc ctccctggtga aaatgaacac ccggtcagtg accaccatgg 540
cggtcattaa cttgggtggg gtccacagcg ttttctgct gacagtgcga tttcgcttga 600
cctacctcat caagaagact tggatgtttg ggctgccctt ctgcaaatat gtgagtgcga 660
tgctgcacat ccacatgtac ctacagttcc tattctatgt ggtgatcctg gtcaccagat 720
acctcatctt cttcaagtgc aaagacaaag tggaattcta cagaaaactg catgctgtgg 780
ctgccagtgc tggcatgtgg acgctggtga ttgtcattgt ggtacccctg gttgtctccc 840
ggtatggaat ccatgaggaa tacaatgagg agcactgttt taaatttcac aaagagcttg 900
cttacacata tgtgaaaatc atcaactata tgatagtcac ttttgtcata gccgttgctg 960
tgattctgtt ggtcttcag gtcttcatca ttatgttgat ggtgcagaag ctacgccact 1020
ctttactatc ccaccaggag ttctgggctc agctgaaaaa cctatttttt ataggggtca 1080
tccttgtttg tttccttccc taccagtctt ttaggatcta ttacttgaat gttgtgacgc 1140
attccaatgc ctgtaacagc aaggttgcat tttataacga aatcttcttg agtgtaacag 1200
caattagctg ctatgatttg cttctctttg tctttggggg aagccattgg ttttaagcaa 1260
agataatggc ttatggaatt gtgtttgtgc cgttagccac aaactacagt attcatattt 1320
gcttccttta tattgggaat aaaaatgggt ataggggagg taagaatggt atttcattac 1380
ttgatcaaaag catgccttga tgtaaccaaa acaaaaggac tataaatgca agagccctca 1440
ttgtagtctt attgggatcc tccatctcga gtgatggcgt acaagacccg tgttgtcgc 1499
```

&lt;210&gt; 23

&lt;211&gt; 2455

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 55012817CB1